SAND 2011-5442C

Challenges for Assessments of Modeling and Simulation in the Large

Greg Weirs
Sandia National Laboratories
P.O. Box 5800, Albuquerque, NM 87185
vgweirs@sandia.gov

Institute for Computing in Science
Verification, Validation, and Uncertainty Quantification Across Disciplines Workshop
August 7-12, 2011

National Nuclear Security Administration

Sandia National Laboratories is a multi-program laboratory managed and operated by Sandia Corporation, a wholly owned subsidiary of Lockheed Martin Corporation, for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-AC04-94AL85000.

Sandia National Laboratories

Who I Am

Greg Weirs

- Academic background: aerospace engineering CFD
- >10 yrs experience in academia & national labs
- Methods and techniques: code development, numerical method development, analyst
- Application areas: Atmospheric entry, combustion, astrophysics, solid mechanics, MHD, V&V, UQ, SA, scientific visualization
- *Clearly, a jack of several trades.*

Outline Of Ideas

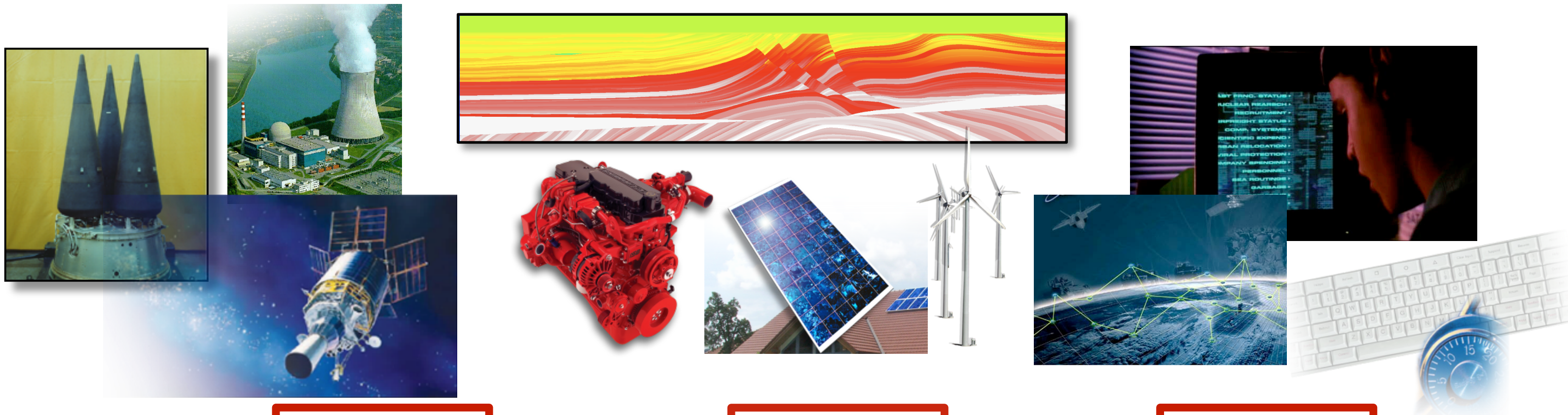
There remain many technical challenges in verification, validation, UQ, etc. BUT:

Many barriers to assessing numerical simulations are non-technical

- Context is everything
- The sociology of the ecosystem or, the ecology of the social system. (?)[1]
- Be skeptical

[1] Sorry, social sciences and humanities are not among my trades…

Sandia occupies a national position in computational science supporting predictive modeling and simulation.



Sandia Strategic Thrusts
- Nuclear Security
- Energy Security
- Cyber Security

Computer Science Application Themes
- Nuclear Weapons Engineering
- Energy/Climate Security
- Cyber Assurance
- *Decision Support*

Computer Science Technology Themes
- Predictive Simulation
- Scalable Computing
- Scalable Informatics
- Cognitive Science
- *Enabling Technologies*

In M&S, you don't know how good (or bad) you are if you don't ask.

- "Due diligence" means asking all the questions, even if you don't think you'll like the answers.

Credit: M. Pilch.

Scientific Software Context

What makes engineering physics modeling and simulation software different?

- ➢ Our simulations provide *approximate* solutions to problems for which we do not know the exact solution.

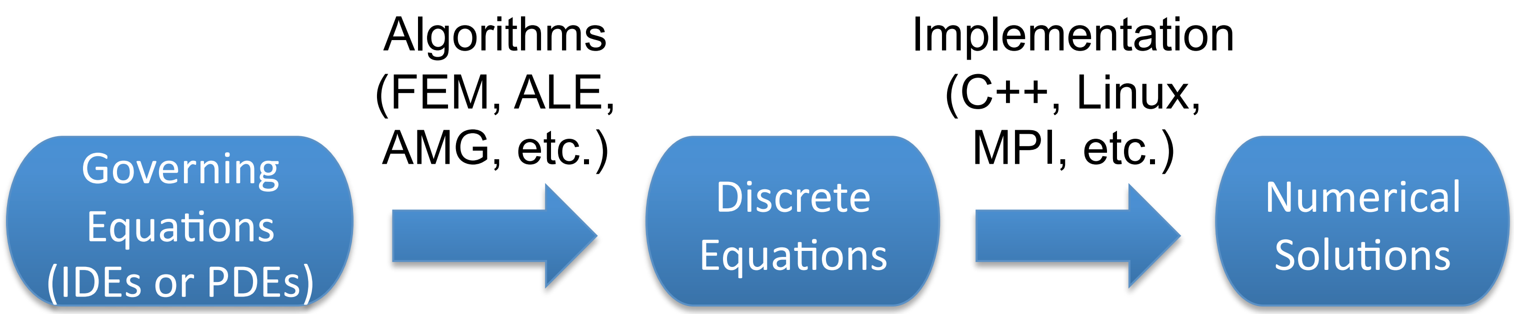
This leads to two more questions:

- How good are the approximations?
- How do you test the software?

A Numerical Simulation is the Conclusion of a Long Development Process

Model: Governing Equations, ICs, BCs, Submodels (constitutive models, closure relations, etc.)
Here, model does *not* mean code

Implementation: Compute the approximate solution



Algorithms: Generate a solvable discrete system; solution of the discrete system is an approximate solution of the model

Different Assessment Techniques for Different Sources of Error

Problem:	Assessment:
• Model(s) not good enough	• Validation
• Numerics not good enough	
• Algorithm is not implemented correctly	• Code verification
• Algorithm is flawed	• Code verification
• Problem definition not good enough	• Uncertainty quantification

Sandia National Laboratories

Putting It All Together

- Code verification, validation, calculation verification, and especially UQ, SA, and calibration -- are tools that can have different uses for different problems

- For the assessment of numerical simulations, a framework of these tools is organized by the different types of things that can go wrong.

  Q:Why?

  A: Knowing what goes wrong is the first step to correcting it.

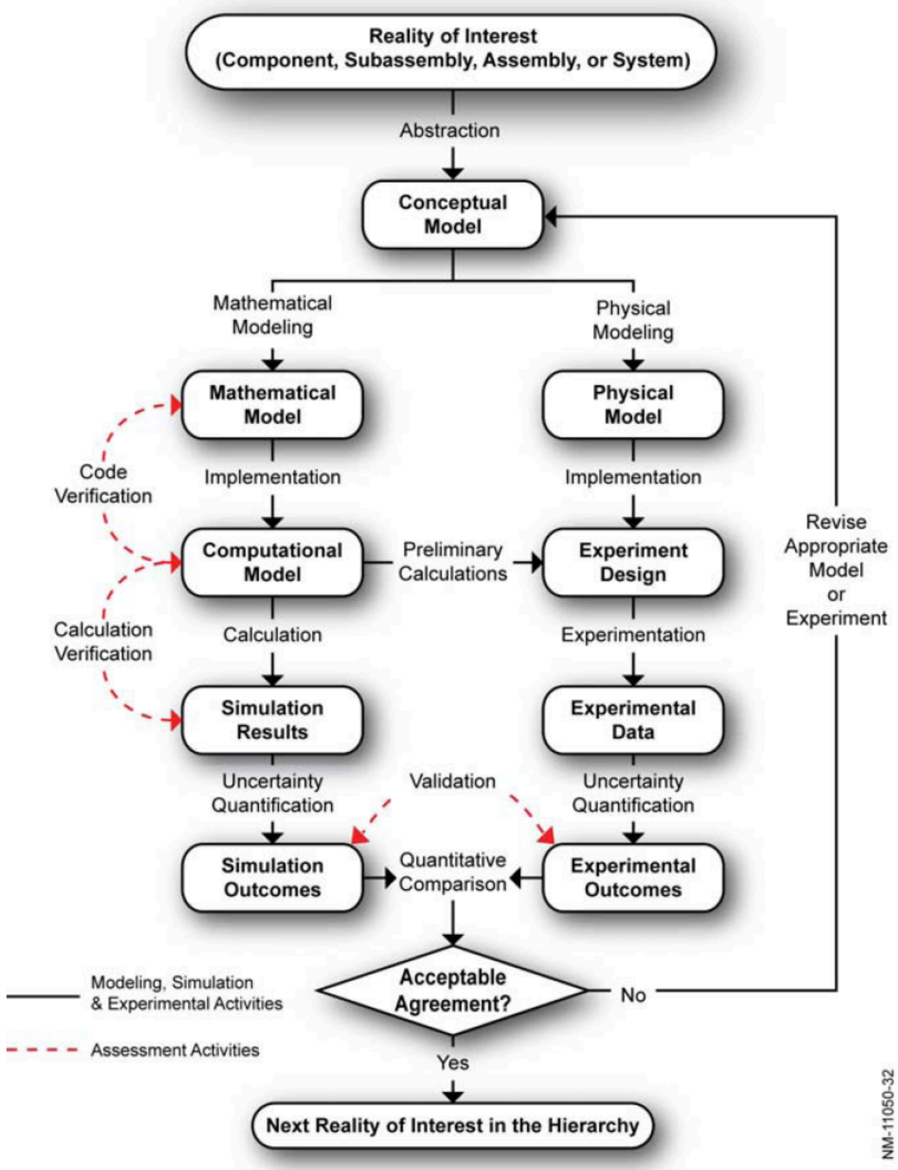
Making a Case Without Gaps

Software Quality Engineering (SQE)	Manage software complexity
Code Verification	Assess algorithms and their implementation vs. exact solutions
Solution Verification	Estimate discretization error
Validation	Assess physics models vs. experimental data
SA / UQ	Assess sensitivity or uncertainty of answer to input parameters

An ingredients list for predictive simulation, *not a menu*.

Sandia National Laboratories

The ASME proposed a V&V workflow.



- This V&V process flowchart is taken from the ASME Solid Mechanics V&V guide.
- Note the positions of code verification, calculation verification, validation, and UQ in this workflow diagram.

ASME, *V&V 10-2006 Guide for Verification and Validation in Computational Solid Mechanics*, American Society of Mechanical Engineers (2006).

Code Verification As A Continuous Process

- To set up a verification problem once takes significant effort – steep learning curve, infrastructure is not in place

```
while 1:
   …
   run verification_suite
```

- Running a verification analysis you have maintained takes minimal work
- Without regular, automated verification testing, verification results go stale quickly - they do not reflect the current state of the code

Code Verification Is Not Free

Principal Costs:

- Infrastructure development
- Test development

Recurring Costs – A tax on development:

- Maintenance of existing tests
- Code development becomes a very deliberate process

Sustainable verification: Benefits outweigh costs

Code Verification Identifies Algorithmic Weaknesses

One purpose of code verification is to find bugs.

- Code verification often finds bugs that are subtle and otherwise difficult to identify.
- The eyeball norm finds most obvious bugs quickly.

Perhaps a better use of code verification is to guide code development.

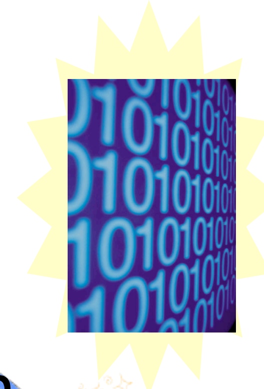
- Some bugs are algorithmic and conceptual.
- Code verification identifies algorithmic weaknesses.
- Large errors are a weakness.

Sandia National Laboratories

The Most Efficient Code Verification is Done by Code Developers

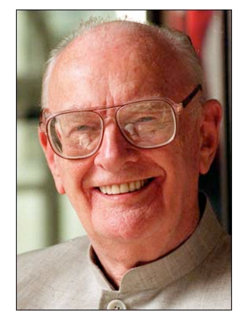
- Code developers best understand the numerical methods they are using
- Code developers are best able to use the results of code verification (and other forms of assessment) to improve the algorithms they use
- Code verification as an accreditation exercise has no lasting impact on code quality

The nature of the code development is a key aspect to consider.



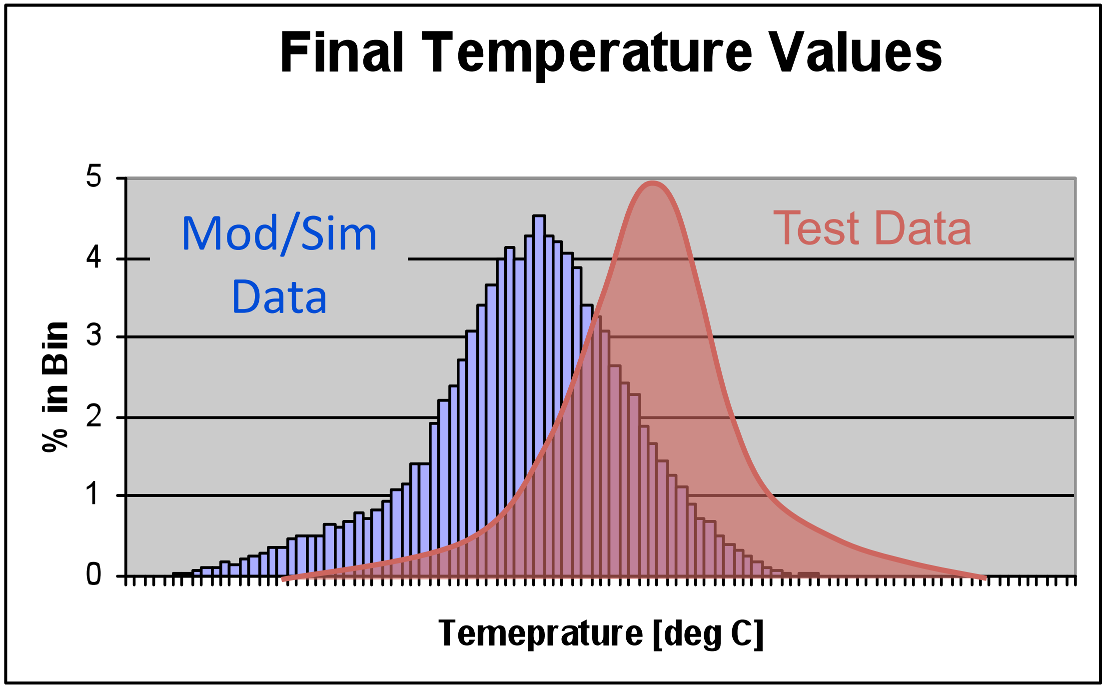
- How well do the code developers understand what they are working on.
- In some cases the key developers have moved on and are not available…

- … leading to the "magic" code issue,
  - "Any sufficiently advanced technology is indistinguishable from magic." Arthur C. Clarke [Clarke's Third Law]
  - Understanding problems can be nearly impossible, or prone to substantial errors,
  - Fixing problems become problematic (bad choices are often made!) as a consequence.



Sandia National Laboratories

A notional example of validation analysis illustrates the incorporation of uncertainty.



Validation:

- Compare simulation data histogram to a test data histogram.
- Quantify amount of “overlap” between histograms.
- Assess sufficiency of overlap.

Uncertainty Quantification:

- UQ methods generate an ensemble of mod/sim data.

- UQ methods are used to generate statistical information on the code output.
  - Probability distribution on Temperature, *given various* $x_1,...,x_N$ *inputs.*
  - Correlations (i.e., trends) of Temperature vs. $x_1,...,x_N$.
  - Mean(T), StdDev(T), Probability($T > T_{critical}$)

Credit: Tim Trucano

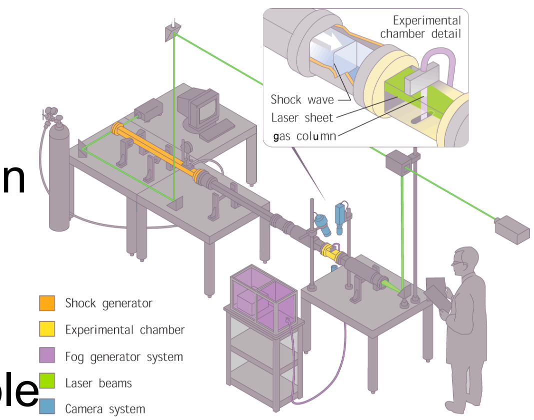
Sandia National Laboratories

Experiments are an inherent element of *any* validation analysis.

- There are different types of experiments:

NTS

LANL shock tube lab



Legacy	vs. “Live”
Performed in the past	Currently undertaken
Often <u>un</u>repeatable	Hopefully repeatable
Limited error information	More error information

Discovery	vs. Validation
Maybe repeatable	Necessarily repeatable
Usual experimental controls	Careful experimental controls
Usual error information	Extensive error information

- Analysts and experimentalists need to interact!
  - The whole *really is* greater than the sum of the parts.
  - You *really do* learn from each other.

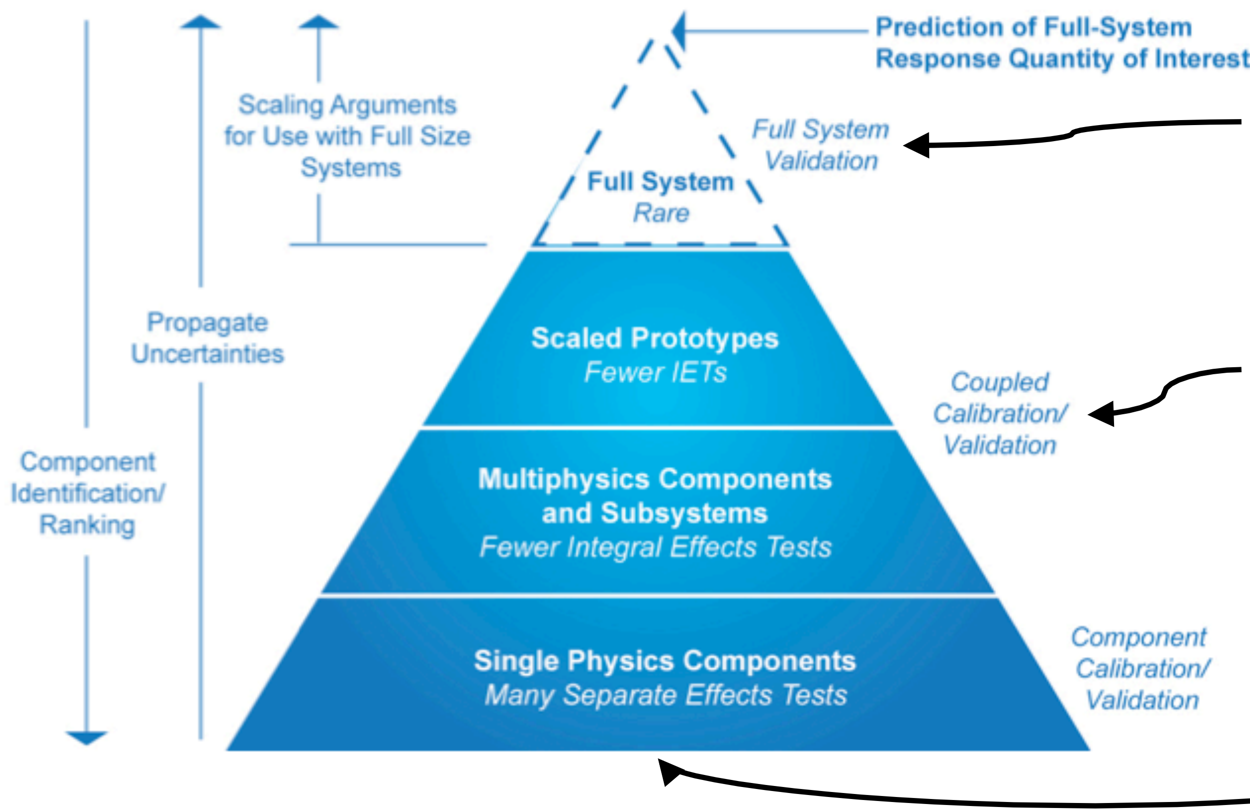
See: W. Oberkampf, “What are Validation Experiments?” *Experimental* Techniques, **25**, pp. 35–40 (2001).

Barriers to synergy between simulations and experiments

- (Project) timing and (project) timescales
- Mutual suspicion of participants
- The language barrier
- Alignment of incentives

These barriers can be overcome, with a lot of work.

Complex problems require a hierarchical approach to validation.



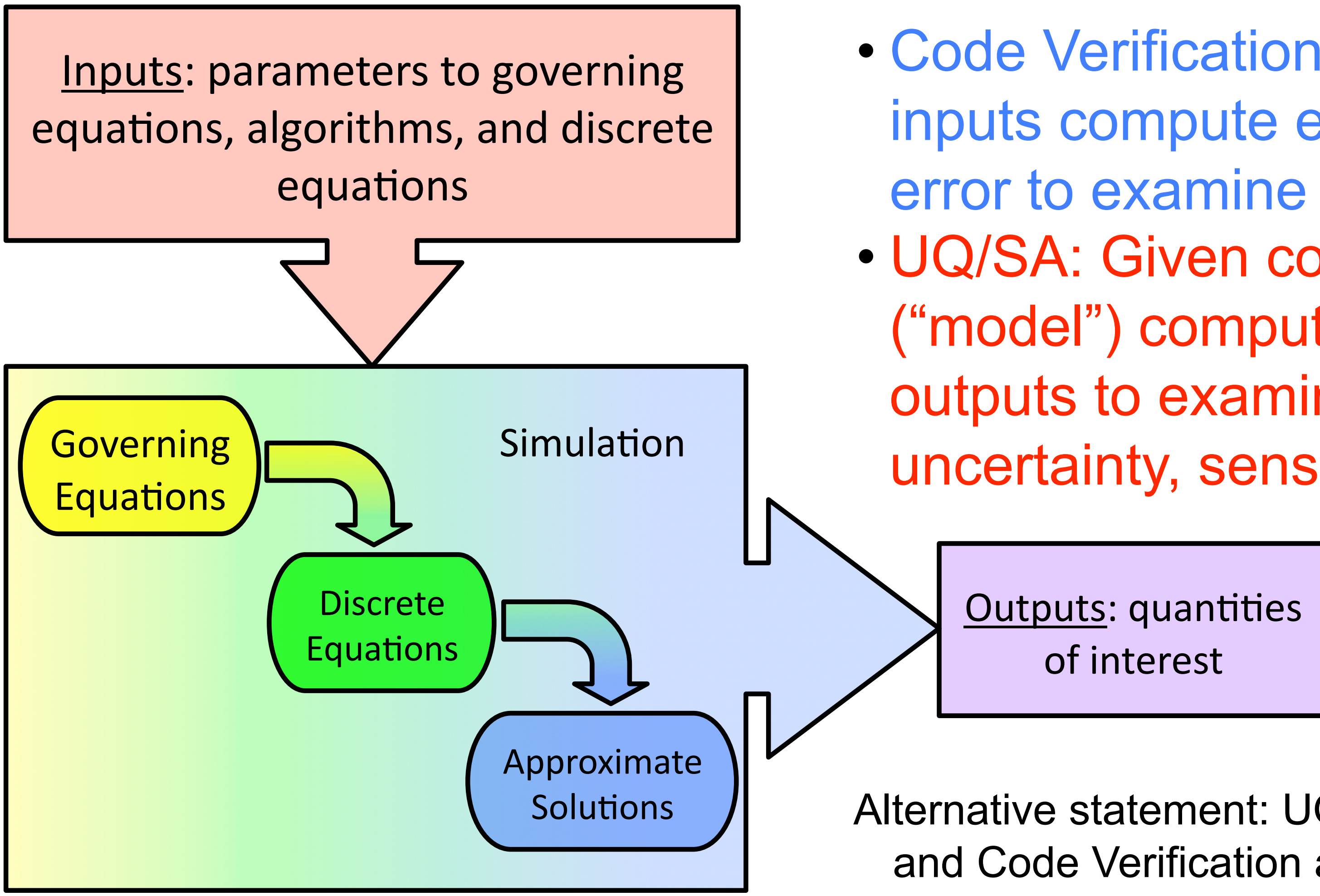
Validation of "full-physics" systems is usually rare.

Validation of "coupled-physics" systems is often complex and challenging.

Validation of "single-physics" systems is the most common analysis.

Credit: D. Kothe et al., Consortium for Advanced Simulation of Light Water Reactors (CASL), 2010.

Verification is orthogonal to UQ and SA



- Code Verification: Given inputs compute exact error to examine code.
- UQ/SA: Given code ("model") compute outputs to examine uncertainty, sensitivity.

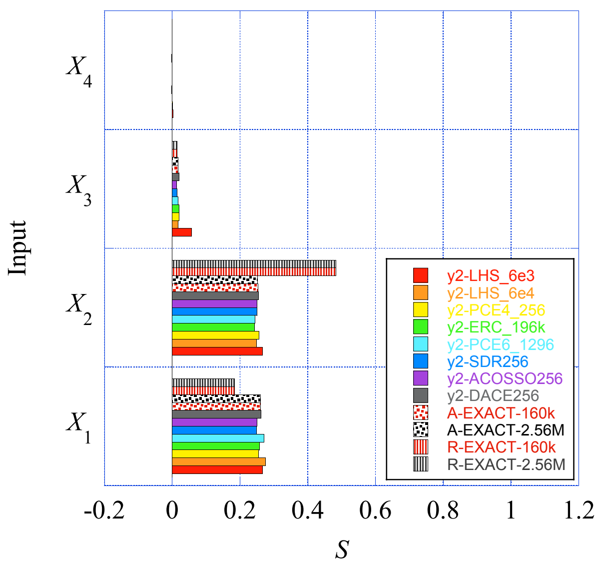
Alternative statement: UQ/SA and Code Verification are complementary.

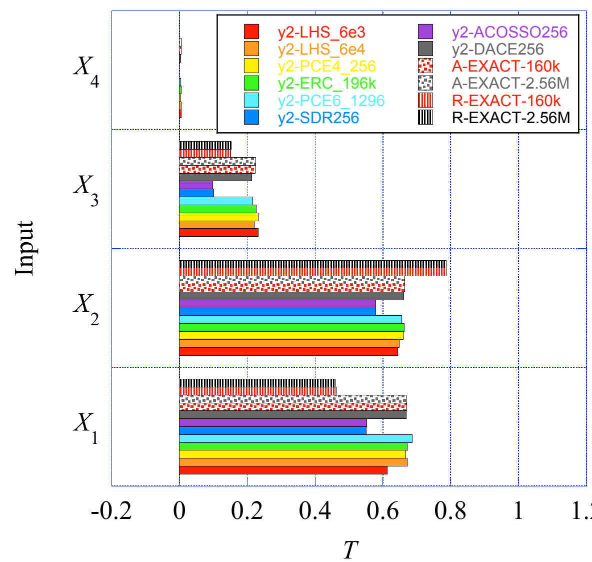
Sandia National Laboratories

Limits of UQ and SA

- Uncertainty quantification does not say anything about the quality of the numerical method.

- Uncertainty quantification does not say anything about the quality of the model.





Sandia National Laboratories

Generalized Polynomial Chaos Expansions

Approximate response w/ spectral projection using orthogonal polynomial basis fns

i.e. $\boxed{R = \sum_{j=0}^{P} \alpha_j \Psi_j(\xi)}$ using

$$\begin{array}{lllll} \Psi_0(\xi) & = & \psi_0(\xi_1)\,\psi_0(\xi_2) & = & 1 \\ \Psi_1(\xi) & = & \psi_1(\xi_1)\,\psi_0(\xi_2) & = & \xi_1 \end{array}$$

This information is very difficult to obtain in the large

- Nonintrusive: es
(regression), tenso
integration)

Wiener-Askey Ge alized F

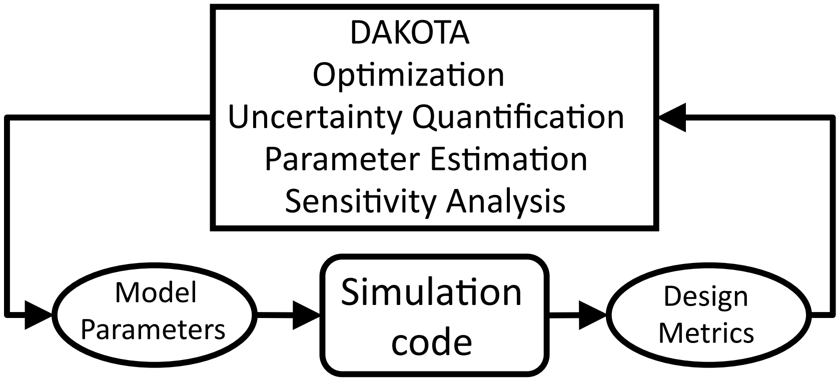
- Tailor basis: o al basis selection leads to expon conv rates

Distribution	ensity function	Polynomial	Weight function	Support range
Normal	$\frac{1}{\sqrt{2\pi}} e^{\frac{-x^2}{2}}$	Hermite $He_n(x)$	$e^{\frac{-x^2}{2}}$	$[-\infty, \infty]$
Uniform	$\frac{1}{2}$	Legendre $P_n(x)$	1	$[-1, 1]$
Beta	$\frac{(1-x)^{\alpha}(1+x)^{\beta}}{+\beta+1 B(\alpha+1,\beta+1)}$	Jacobi $P_n^{(\alpha,\beta)}(x)$	$(1-x)^{\alpha}(1+x)^{\beta}$	$[-1, 1]$
Exponential	e^{-x}	Laguerre $L_n(x)$	e^{-x}	$[0, \infty]$
Gamma	$\frac{x^{\alpha} e^{-x}}{\Gamma(\alpha+1)}$	Generalized Laguerre $L_n^{(\alpha)}(x)$	$x^{\alpha} e^{-x}$	$[0, \infty]$

Credit: M. Eldred, L. Swiler

andia National Laboratories

Sandia's toolkit DAKOTA is software for optimization and UQ .



DAKOTA is openly available:

http://dakota.sandia.gov/

- Design and Analysis toolKit for Optimization and Terascale Applications includes a wide array of algorithms for UQ, SA, and optimization.
- Other packages share some similar functionality:
  - Minitab statistics package
  - JMP statistical software
  - Mathematica
  - Matlab with Statistics Toolbox
  - R or S+ language
  - Simlab
  - Excel add-ins, such as @Risk and Crystal Ball

Sandia National Laboratories

DAKOTA's new user challenge

You are tasked with using DAKOTA to find the uncertainty of figure of merit Y given uncertain parameters X. You know how to run your (sophisticated) code to get Y from X.

+ Training classes are offered several times per year
+ There are about 1500 pages of documentation and a wide variety of example problems to download and play with.
+ There is a user's list and a developer's list, and the code team is responsive.

Barriers to user adoption can run deep

So what's the problem?

- Your have to learn how to specify DAKOTA's input deck. (Classes and a new GUI help here.)
- You have to hook up your code to DAKOTA. Scripts are required to:
  - Have DAKOTA specify the inputs to your code
  - Extract Y from your code or output files and return it to DAKOTA.
- You don't know anything about UQ.

User education is the biggest hurdle

V&V Is a Tough Sell.

V&V is expected, but not well understood, by decision makers.

- V&V is, in a nutshell, all about putting "correct" math methods and physics models in our codes.
- We're <u>expected</u> to produce "correct" codes.
- "If you haven't been doing V&V all along, then what have you been doing with my _____ money?"

…but the importance of V&V is increasing.

What's different now?

- Computational simulation is different now than 10-20-30 years ago (e.g., auto industry, aircraft industry, nuclear weapons industry)
  - We're making million/billion $ decisions that are heavily influenced by comp. sim.
- Definition of "correct codes/models" (see previous) is now changing.
- "Before I spend $M/$B on a decision, I want evidence of the correctness of your simulation model and results."

Using simulation results to aid decision making sells.

- Decision making is based on knowing the tradeoffs for competing objectives, due to variations in designer-controllable parameters.
- Quantities of interest: cost & performance
- This sells (re: facility design hardness study):

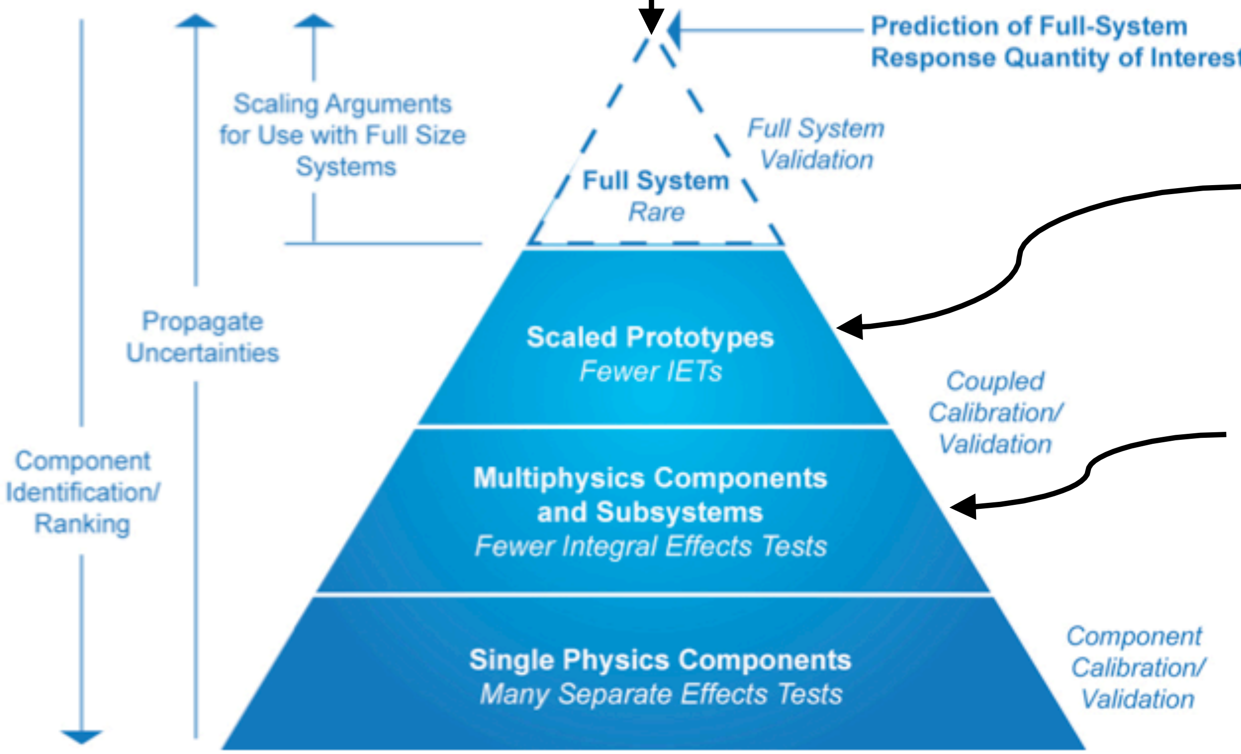
  "If you increase factor1 by A% and lower factor2 by B%, you reduce cost by X% and decrease the probability of kill by Y %."

  "By the way, here is the evidence (tucked away in a report appendix) for the validity of predictions A, B, X, and Y."
- This also sells:
  - "If were going to perform a comp. sim. study that influences a $M/$B decision, then let's carve out $m to run a V&V study to make sure we're getting good data, and $n to perform an adequate sensitivity/uncertainty analysis."

The key is how results are communicated.

- V&V doesn't sell for it's own sake.
  - Decision makers don't care about the rate of convergence of an iterative mathematical method, or % line coverage of tests.
  - For $M/$B issues, decision makers do care that you got the right answer and they expect a technical pedigree (aka "provenance") for your work.
- V&V sells when it is included as an aid to decision making.
  - i.e., when V&V provides supporting evidence (provenance) to sensitivity analysis and UQ results on relevant technical/financial issues.

The fundamental tension: find out as much as you can, but recognize you can't eliminate all risk



If we could validate here, we wouldn't need simulations.

Is there any evidence for the possibility of a catastrophic failure?

Assessments tell where to focus resources to reduce uncertainties.

We can often make predictions with confidence (but it takes a lot more work than you might think.)

Image: Kothe et al., Consortium for Advanced Simulation of Light Water Reactors (CASL) proposal, 2010.

Sandia National Laboratories

Tim Trucano's observations on V&V…

- Key V&V themes have not changed "for decades":
  - "Codes are not solutions, people are solutions."
  - "Credibility of computational simulations for defined applications is evolutionary..."
  - "… at worst, credibility is non-existent in specific applications."
  - "Single calculations will never be 'the right answer' for hard problems."
  - "Real V&V and real UQ are a lot of work."

Old Ideas In V&V

Tim Trucano
Optimization and Uncertainty Quantification (01441)
tgtruca@sandia.gov

August, 2011

Too many acknowledgments are appropriate to list adequately, but I will emphasize my extreme debt to Marty Pilch and Bill Oberkampf

3 August 2011 Old Ideas in V&V Page 1

Sandia National Laboratories is a multi-program laboratory managed and operated by Sandia Corporation, a wholly owned subsidiary of Lockheed Martin Corporation, for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-AC04-94AL85000.

- Trucano's four insights on V&V:
  1. "V&V — pay me now or pay me later."
  2. "Journal editorial policies and practices must change."
  3. "Ask 'What's good enough?'"
  4. "Saying you don't need verification is like saying you don't need oxygen."

Where the Talk Comes From

- This talk contains points of view that have evolved over nearly two decades—and continue to evolve.
- The majority of the ideas in these slides—and many of the actual slides—come from a colleagues at SNL, LANL, LLNL, universities, and institutes, including:

  Brian Adams, Mark Anderson, Ken Alvin, Scott Brandon, Jerry Brock, Hugh Coleman, Scott Doebling, Alireza Doostan, Kevin Dowding, Luís Eça, Mike Eldred, Tony Giunta, John Helton, Dave Higdon, François Hemez, Jan Hesthaven, Rich Hills, Martin Hoekstra, Gianluca Iaccarino, Richard Klein, Patrick Knupp, Sergei Kucherenko, Habib Najm, Bill Oberkampf, Marty Pilch, Marco Ratto, Bill Rider, Pat Roache, Vicente Romero, Chris Roy, Joe Sefcik, Andrea Saltelli, Didier Sornette, Fred Stern, Laura Swiler, Jim Stewart, Stefano Tarantola, Tim Trucano,…
- NNSA (the ASC program) continues to support V&V.

  In particular, Gil Weigand, Dimitri Kusnezov, Bob Meisner.

Sandia National Laboratories